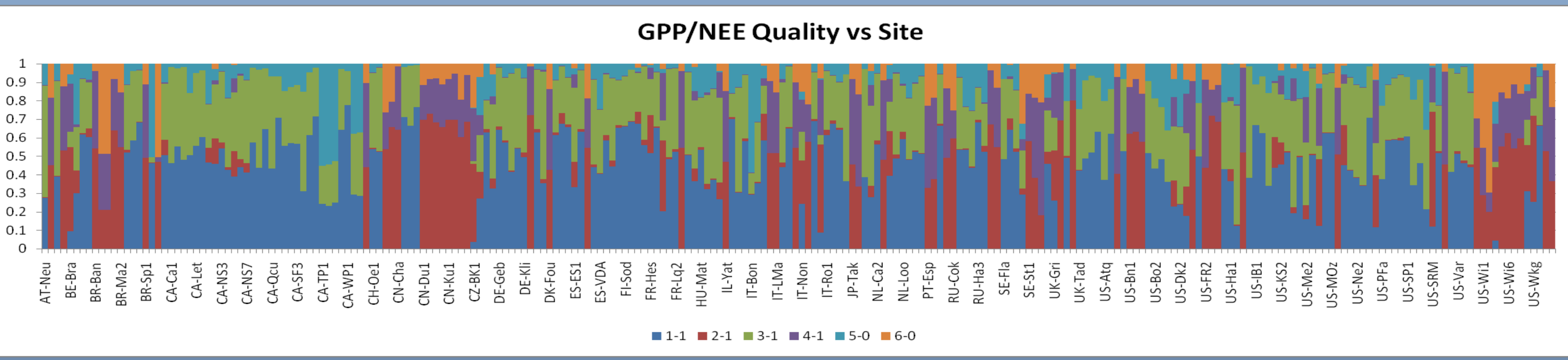
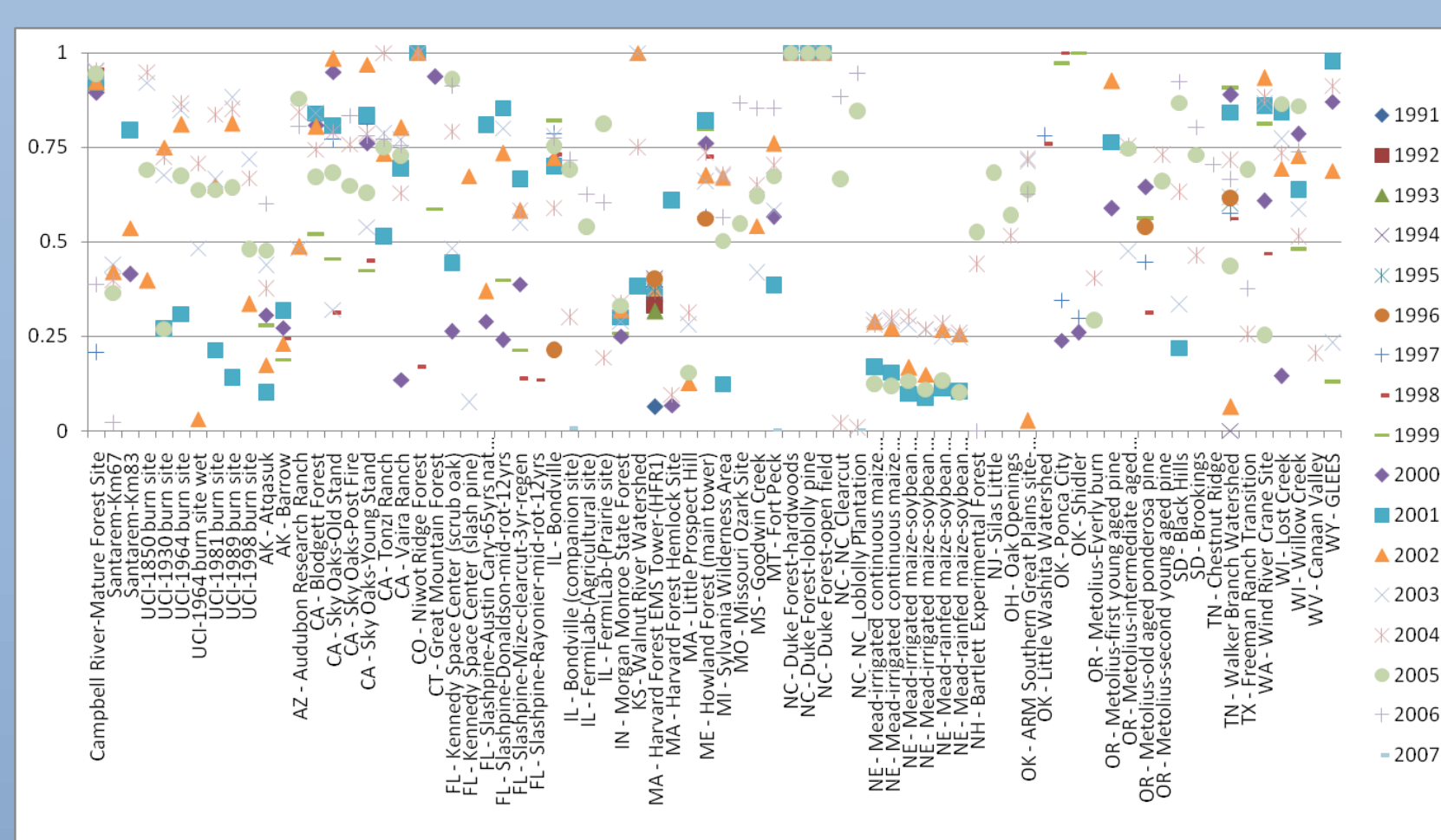
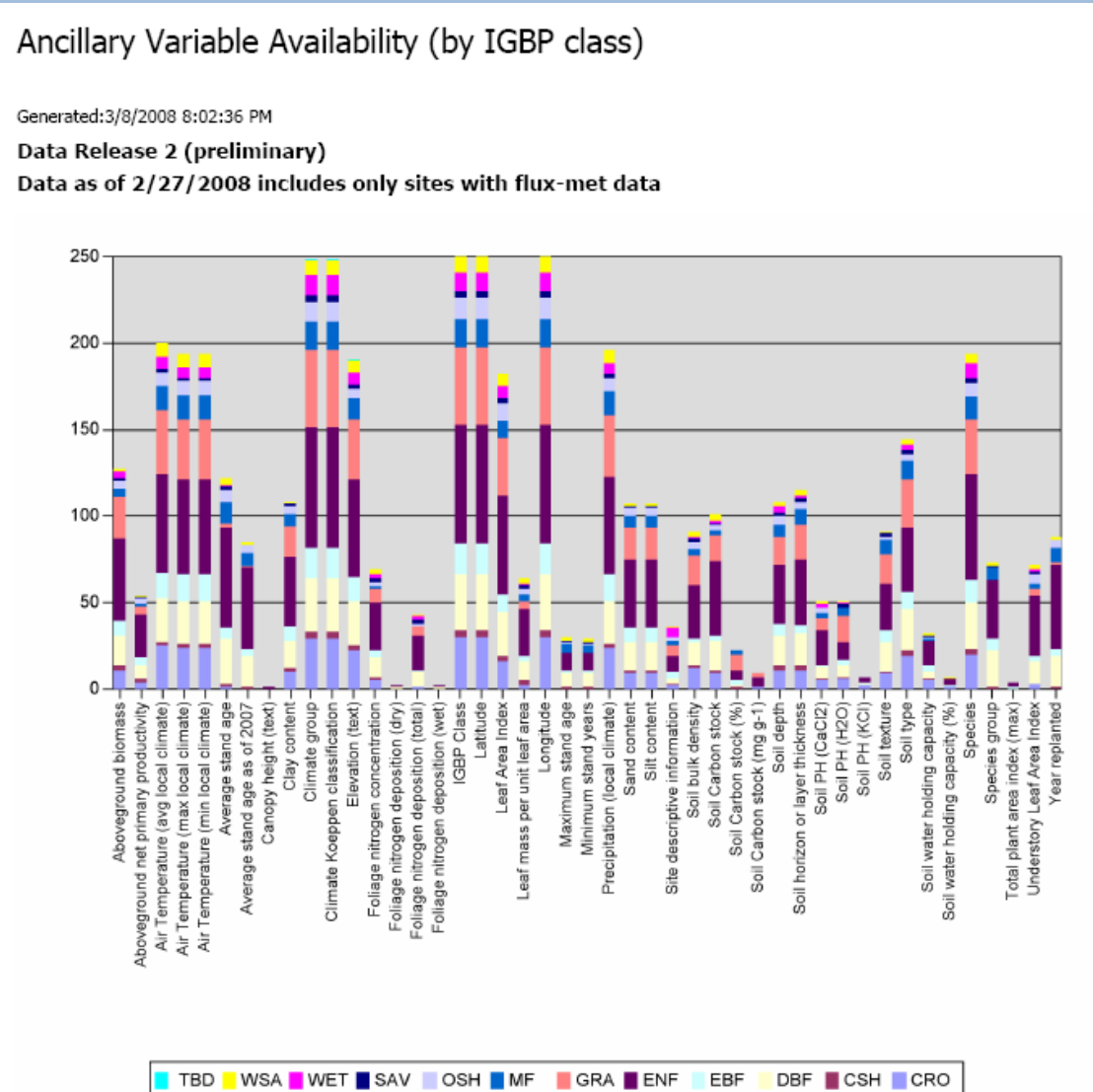


# AN EVOLVING LA THUILE FLUXNET DATASET AND SUPPORT INFRASTRUCTURE

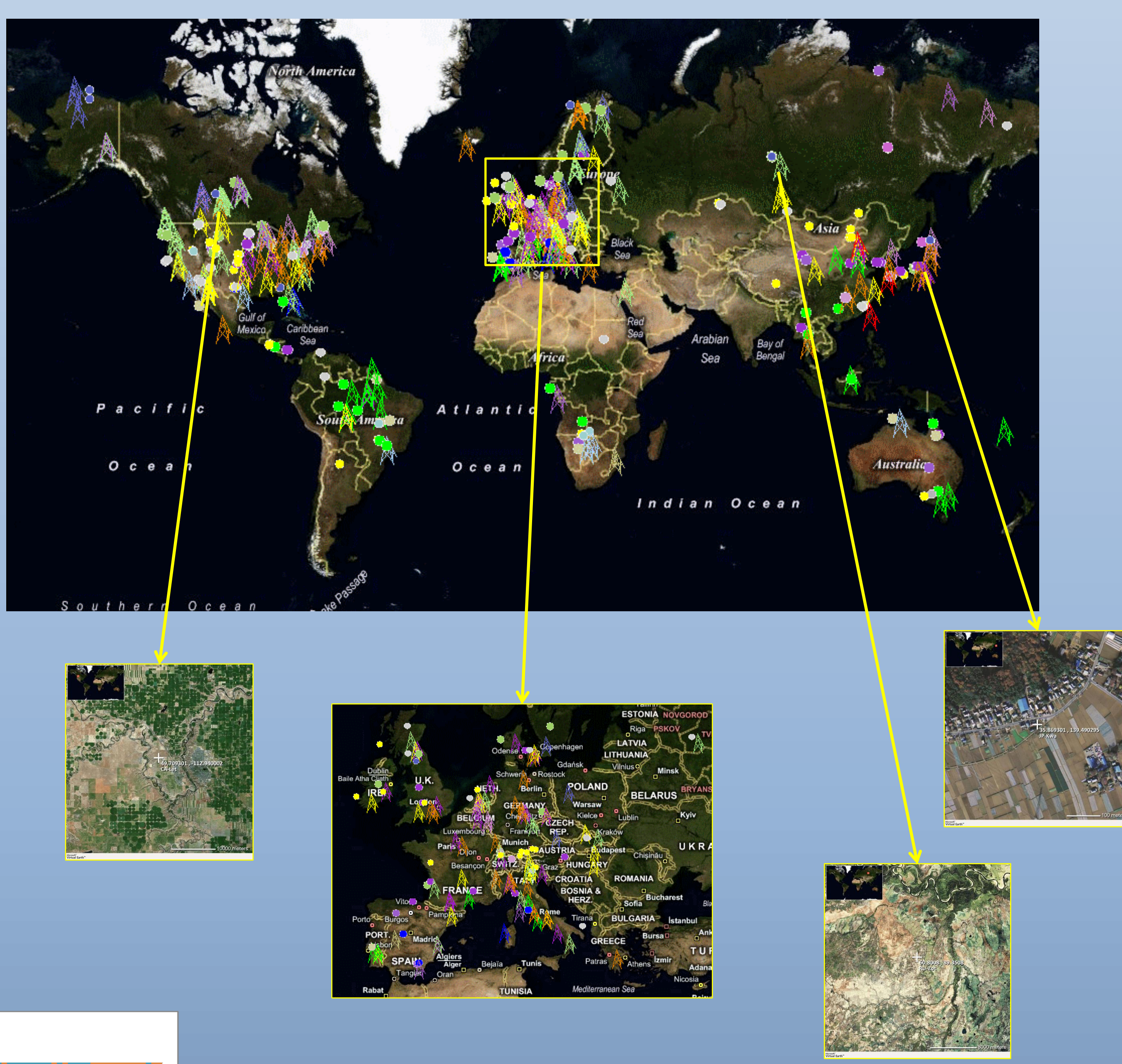
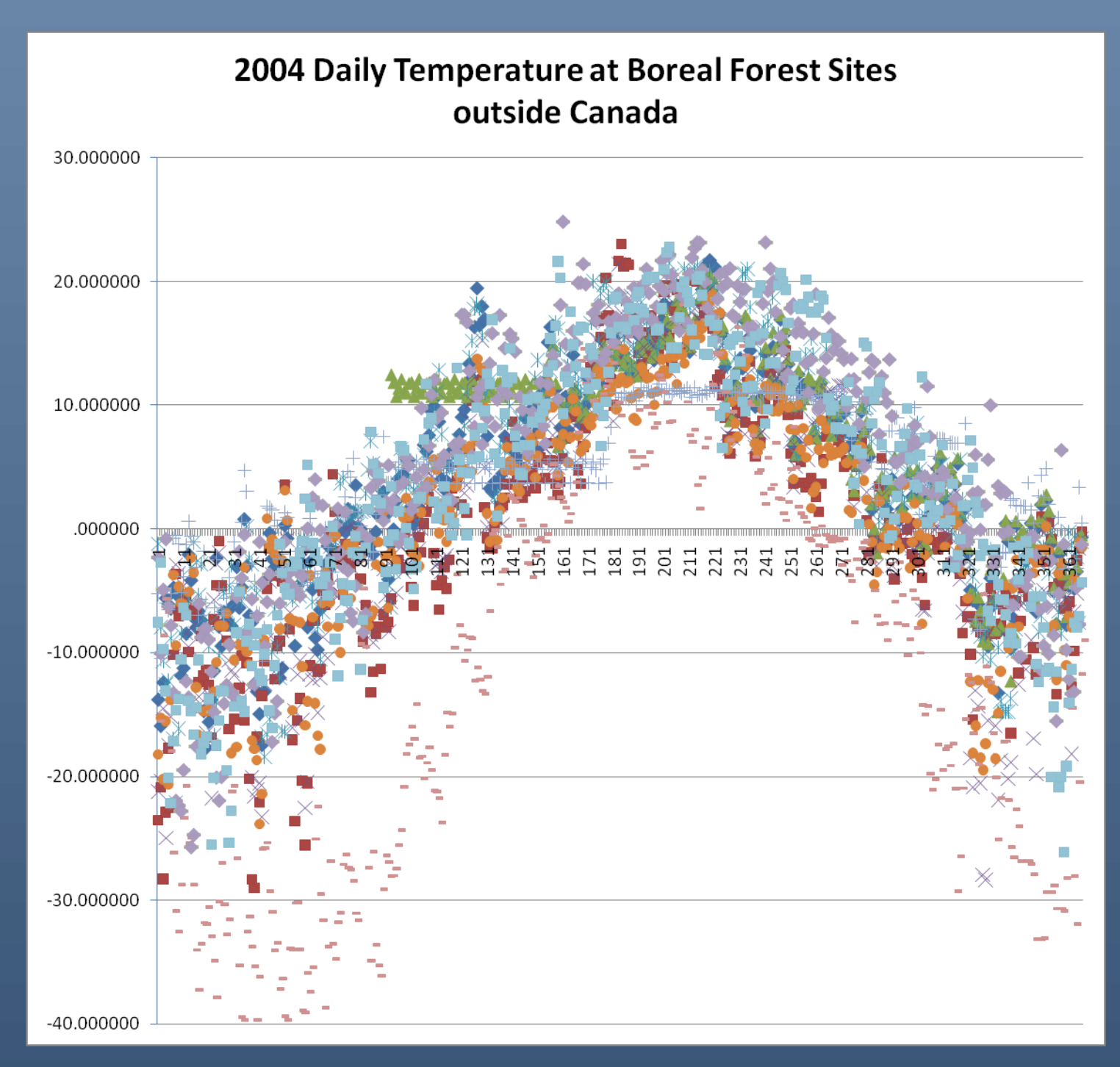
Deb Agarwal<sup>abe</sup> (daagarwal@lbl.gov), Dennis Baldocchi<sup>ae</sup> (baldocchi@nature.berkeley.edu), Monte Goode<sup>abe</sup> (MMGoode@lbl.gov), Marty Humphrey<sup>j</sup> (humphrey@cs.virginia.edu), Catharine van Ingen<sup>d</sup> (vaningen@microsoft.com), Dario Papale<sup>l</sup> (darpap@labecomail.agraria.unitus.it), Markus Reichstein<sup>c</sup> (mreichstein@bgc-jena.mpg.de), Matthew Rodriguez<sup>bf</sup> (mar008@cs.ucsd.edu), Youngryel Ryu<sup>ae</sup> (ryru@nature.berkeley.edu), and Rodrigo Vargas<sup>e</sup> (rvargas@nature.berkeley.edu) Berkeley Water Center<sup>a</sup>; Lawrence Berkeley National Laboratory<sup>b</sup>; Max Planck Institute for Biogeochemistry, Germany<sup>c</sup>; Microsoft Research<sup>d</sup>; University of California, Berkeley<sup>e</sup>; University of California, San Diego<sup>f</sup>; University of Tuscia, Italy<sup>i</sup>; University of Virginia<sup>j</sup>;

## ABSTRACT

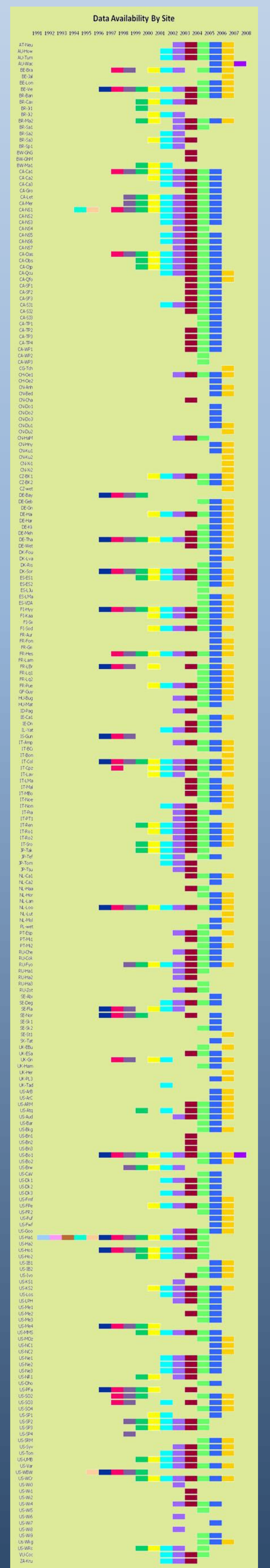
The global FLUXNET synthesis dataset contains over 960 site-years from over 250 sites and continues to grow. This growth reflects previously unavailable processed flux-met data becoming available. The ancillary site information such as biological and disturbance data also continues to evolve. As analysis efforts have begun, issues with the underlying data are often uncovered and the data are corrected when possible. The size of this dataset (90GB) makes finding site years of interest and tracking versions difficult for individual users.



Fluxnet-LatHule									
Dataset Information									
CumulativeValuesByVegType									
Excel Web Access - CumulativeValuesByVegType									
Working Group: All Current Fluxnet Daily									
Note: Cumulative annual values are computed by averaging all reported yearly calculations including partial years and all quality values (0<qOK<1).									
Avg(qOK) > .9 and Avg(count) <= 330									
Avg(qOK) < .9 and Avg(count) >= 330									
no qOK and Avg(count) >= 330									
no qOK and Avg(count) < 330									
Precip									
Rd									
Rq									
Rn									
SWC									
Ta									
Ts									
AT-Hu	Austria - Neustift/Stubai Valley	GRA	1,630.2953	6,7924	4,0136	10,3777	25,8592	30,0698	
AU-Ho	Australia - Howard Springs	WSA	1,123.3560	5,5826	2,8324	9,3362	10,4108		
AU-Tu	Australia - Tumburumba	EBF	892.3947	0.6694	3,5614	23,9029	11,3368	10,5327	
AU-Wa	Australia - Wallaby Creek	EBF	1,633.4386	3,2393	1,1563	54,2106	7,8948	9,4101	
BE-Er	Belgium - Brasschaat (De Jager)	MF	653.7283	2,4072	2,1030	10,8849	11,9498		
BE-Ja	Belgium - Jalhay	MF	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BE-Le	Belgium - Lonze	CRO	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BE-Ve	Belgium - Vleesdam	MF	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BR-Ba	Brazil - Ecolite Bananal Island	EBF	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BR-Ca	Brazil - Caxiuan Forest-Almerim	EBF	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BR-Cl	Brazil - Rond - Faz. Nossa Senhora	GRA	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BR-JZ	Brazil - Rond - Rocio Jara 3/ Parana	EBF	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BR-Ma	Brazil - Manaus - FZK K34	EBF	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BR-Sa	Brazil - Santarem-Km67-Primary	EBF	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BR-Sa	Brazil - Santarem-Km77-Pasture	CHO	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BR-Sa	Brazil - Santarem-Km83-Logged	EBF	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BR-Sp	Brazil - Sao Paulo Cerrado	EBF	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BW-Gh	Botswana - Ghanzi Grass Site	SAV	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BW-Gh	Botswana - Ghanzi Mixed Site	WSA	1,003.0673	3,1911	44,9871	26,4727	25,8225		
BW-Ma	Botswana - Maun-Mopane	WSA	1,003.0673	3,1911	44,9871	26,4727	25,8225		
CA-Ca	Canada - British Columbia	ENF	1,003.0673	3,1911	44,9871	26,4727	25,8225		



We have developed a Scientific Data Server which enables browsing of the data online, data download, version tracking, and the simple production of data summary products. The Scientific Data Server allows individual researchers to concentrate on science rather than data management. We leverage database tools such as data cubes and web reports to enable Excel pivot table, MatLab, and browser access to the data. The data cubes provide organization and aggregation of data along dimensions including time such that it is easy to retrieve daily, monthly, and yearly calculated values. In addition, we have leveraged available collaboration technology (SharePoint) to provide public pages describing the dataset, provide proposer and site PI data access, track data versions and updates, notify users of server updates, and enable support for contact between site PIs and researchers using the data.



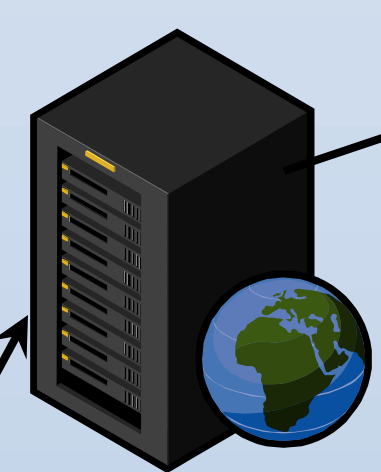


# Challenge is to Connect Data, Resources, and People

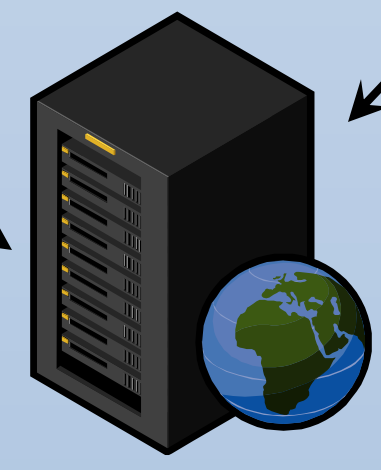
## Distributed Data Sets



## Catalog Database



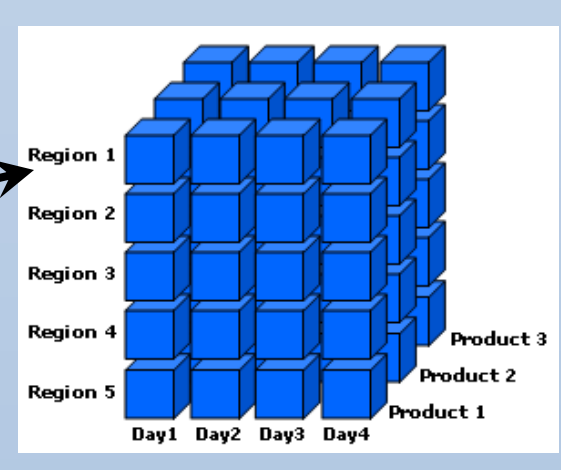
## Archive Databases



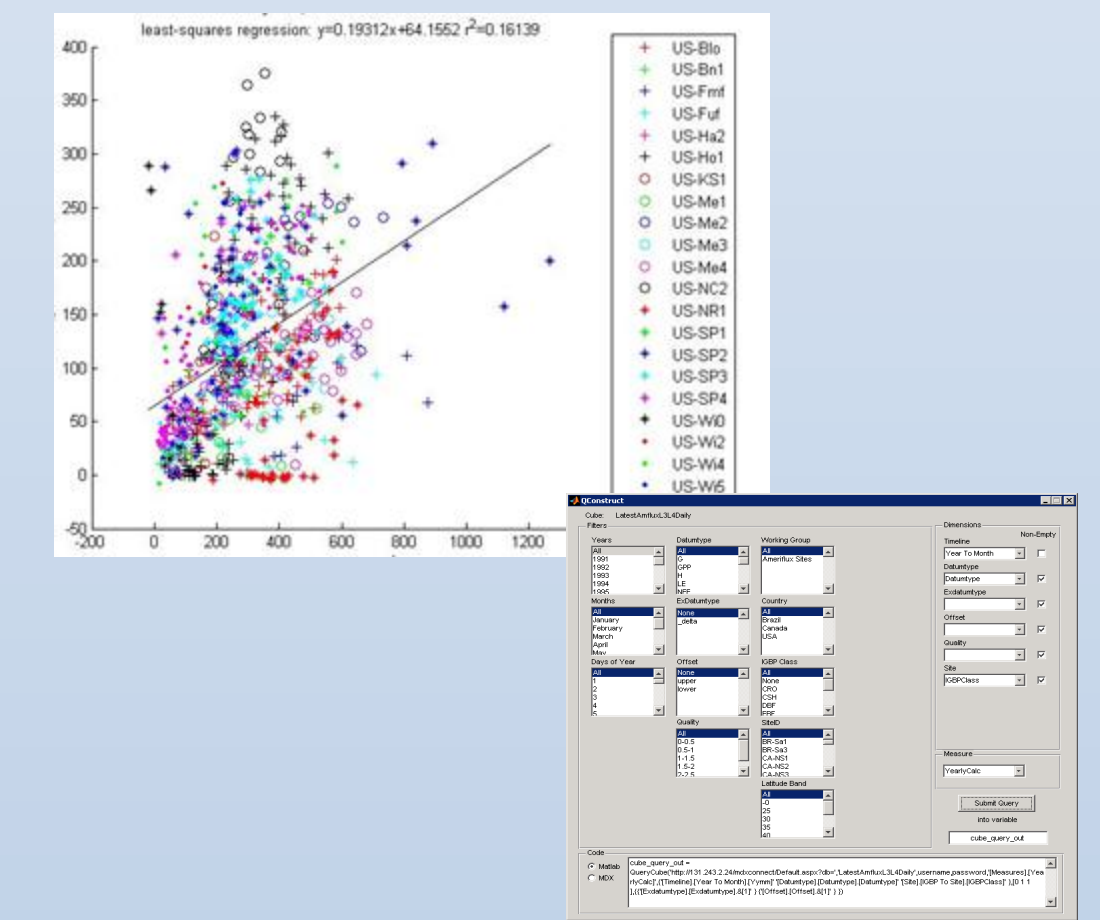
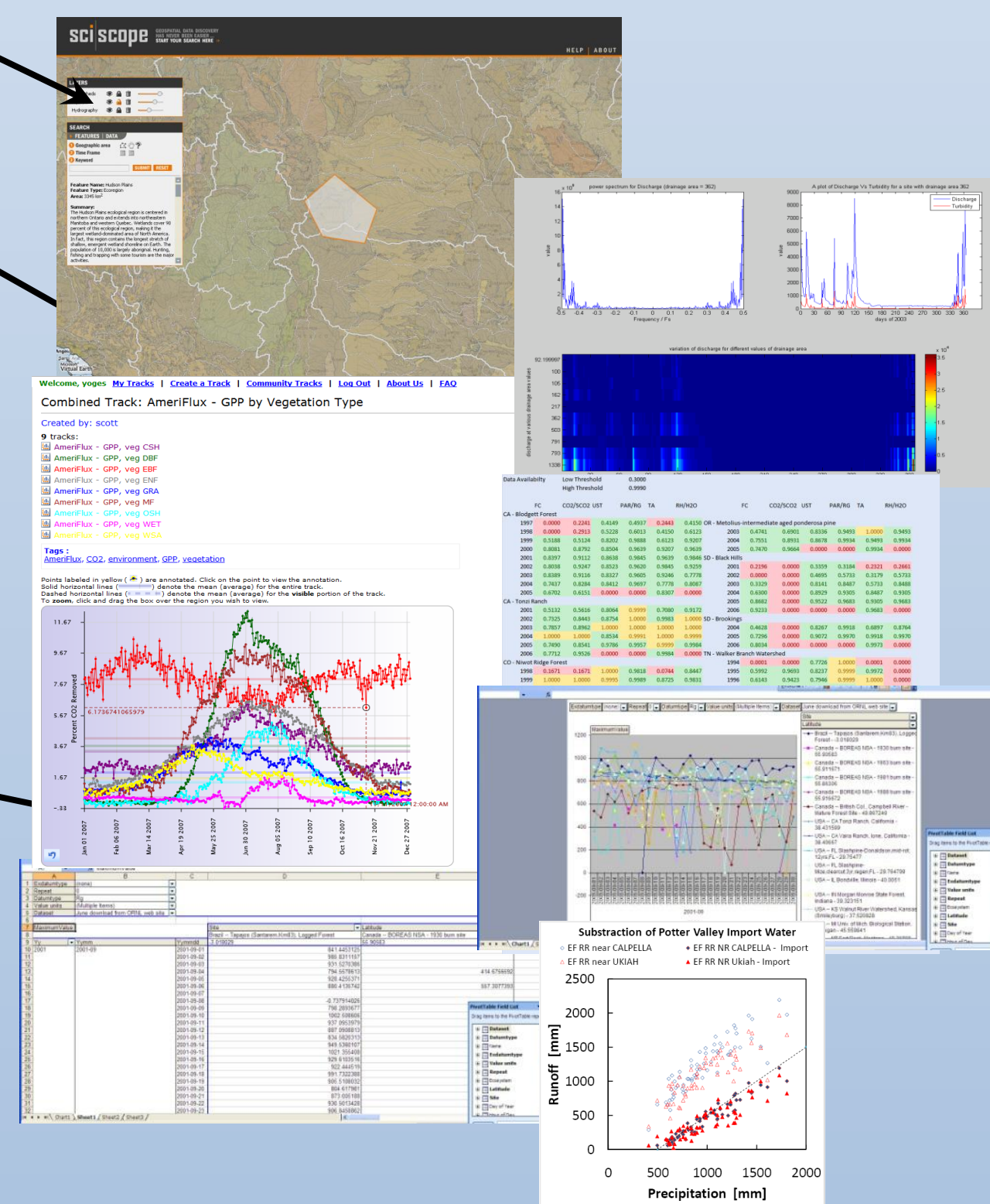
## Models



## Data Cubes



## Visualizations

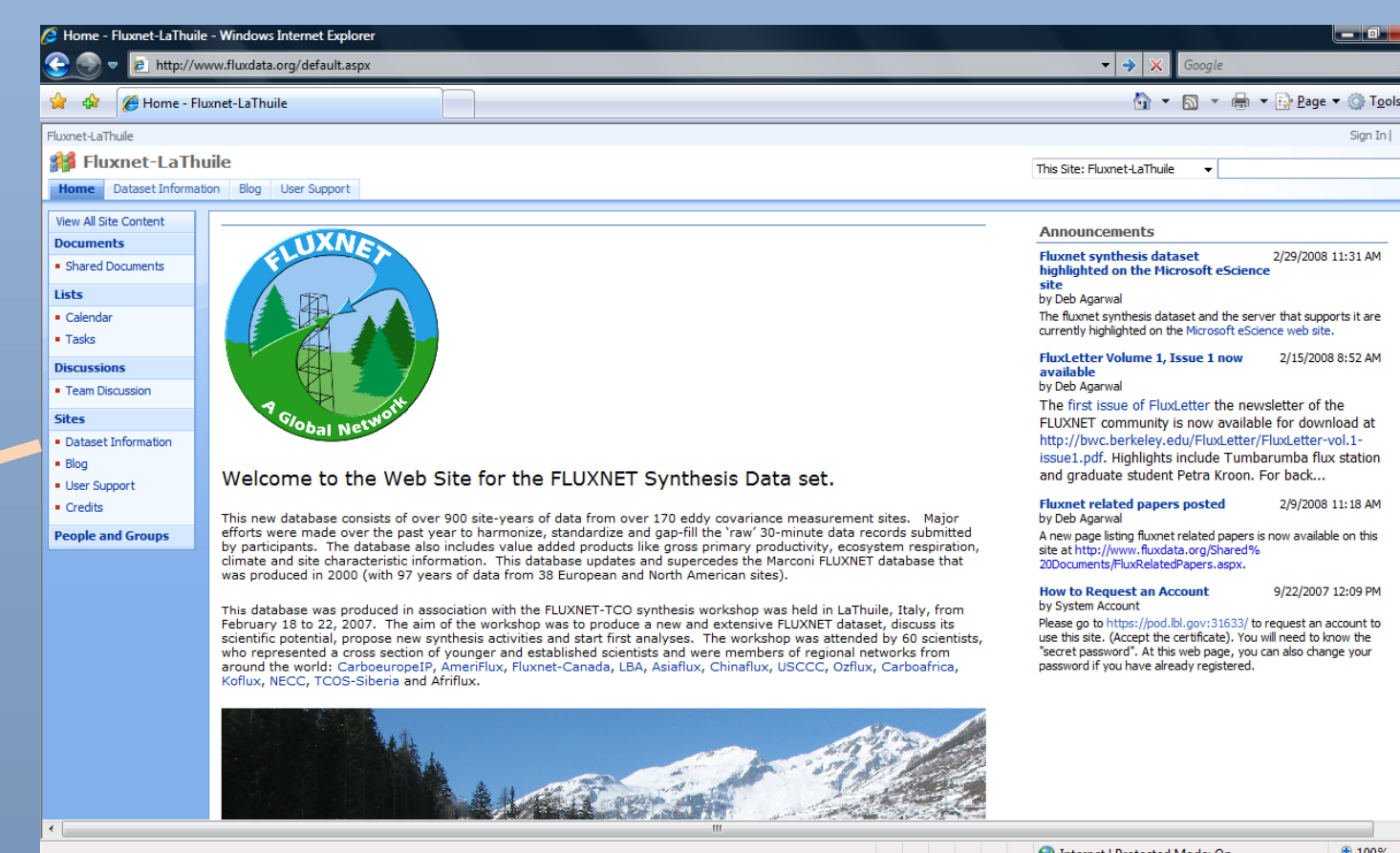


## Published scientific results

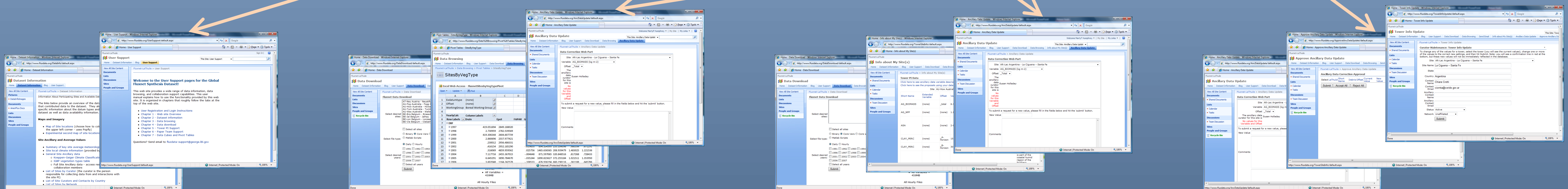
## Fluxnet data server – www.fluxdata.org

- Common data repository capturing provenance information and common data processing
  - Ancillary and biological data incorporated
  - Data browsing, mining, and plotting capabilities
  - Data repository that is easy to maintain, load with data, and expand capabilities
  - Collaboration space to enable interaction between PIs and proposers and among proposal teams
  - Data download and access (protected and tracked)
- Reduce cut & paste manipulation of data  
Dramatically increase the scale of the data that can be analyzed

## Collaborative Interface with Customizable views/functionality based on role(s)



## Powered by Microsoft Office Sharepoint Server 2007



For the public: Dataset and tower information, user support, data summaries, list of synthesis teams, and published results

For the scientists analyzing the data: Download data, interactively explore data, provide ancillary data updates/corrections, and interact with tower owners

For the tower owners: Download curated tower data, review/update info about the tower, provide ancillary data updates/corrections, and interact with synthesis teams

For the data curators: Download data, provide ancillary data updates/corrections, reject/accept proposed data updates/corrections, and synchronize across data sources



<http://www.fluxdata.org/>

For further information e-mail: [fluxdata-support@george.lbl.gov](mailto:fluxdata-support@george.lbl.gov)

